13 En estadística se define la función de densidad normal o de Gauss por la expresión

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

en el intervalo $(-\infty, \infty)$.[2] Demostrar que

(a) $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ (b) $\int_{-\infty}^{+\infty} x f(x)\,dx = \mu$ (c) $\int_{-\infty}^{+\infty} x^2 f(x)\,dx = \sigma^2 + \mu^2$

(*Indicación:* Usar la sustitución $u = (x-\mu)/\sqrt{2}\sigma$, junto con (11.13) y el resultado del Ejemplo 11.9.)

11.4 UNA NOTA SOBRE DISTRIBUCIÓN DE RENTAS Y CURVAS DE LORENZ

Explicamos en la Sección 10.4 que, si $f(r)$ es la función de densidad de la distribución de renta de una población de n individuos, entonces $n\int_a^b f(r)\,dr$ es el número de individuos con rentas en el intervalo $[a, b]$ —véase ecuación (10.27). Además, $n\int_a^b r\,f(r)\,dr$ representa la renta total de esos individuos —véase ecuación (10.28).

La **Curva de Lorenz** es una herramienta estadística que sirve para describir algunas características importantes de cualquier distribución de renta.[3] Esta curva se basa en las cuotas de participación en la renta total que corresponden a los diferentes grupos de individuos de la población, empezando por los más pobres y terminando por los más ricos. Consideremos, por ejemplo, los datos de la Tabla 11.1.[4] La tabla indica que, aparentemente, las desigualdades aumentaron en EE.UU. durante la década de los 80 y disminuyeron en Holanda durante el periodo, mucho más dilatado, de 1959–1985. La distribución en Holanda en 1959 era muy cercana a la de EE.UU. en 1980. La distribución mundial así calculada está próxima a un extremo.

TABLA 11.1 *Cuotas de participación en la renta total*

Grupo de	*EE.UU*		*Holanda*		*Mundial*
renta	*1980*	*1990*	*1959*	*1985*	*1989*
Quinto inferior	*5,2*	*4,6*	*5,0*	*7,8*	*1,4*
Segundo quinto	*11,5*	*10,8*	*11,9*	*13,9*	*1,9*
Tercer quinto	*17,5*	*16,6*	*17,4*	*18,1*	*2,3*
Cuarto quinto	*24,3*	*23,8*	*22,7*	*23,4*	*11,7*
Quinto superior	*41,5*	*44,3*	*43,0*	*36,7*	*82,7*

Se confirman estas intuiciones previas mediante un análisis más cuidadoso basado en las curvas de Lorenz. Para construirlas acumulamos primero las rentas de los diferentes quintos de la población de tal manera que los cinco nuevos grupos que consideramos son respectivamente el 20% inferior, luego el 40% inferior, el 60% inferior, el 80% inferior y, finalmente, la población completa. Esto da

[2] La gráfica de esta función se llama la campana de Gauss por su forma. Está recogida en los billetes de 10 marcos alemanes de 1989, junto con el retrato de su inventor Carl Friedrich Gauss (1777–1855).

[3] El nombre le viene del estadístico norteamericano Max Otto Lorenz, quien la introdujo como uno de los "Métodos para Medir la Concentración de Riqueza" (más bien que renta) en un artículo publicado en el *Journal of the American Statistical Association*, 1905.

[4] Los datos para los EE.UU. están tomados de la oficina del censo. Los datos para Holanda provienen de la Oficina Central de estadística. Los datos mundiales están tomados del programa de la ONU *Human Development Report* para 1992. En verdad no existen datos de esta clase a nivel mundial. Las cifras que se dan representan lo que sería la distribución mundial de la renta si el producto nacional bruto de cada país estuviese perfecta y equitativamente distribuido como renta entre los habitantes de ese país. Sin embargo, no hay razones para pensar que las cifras resultantes exageren la verdadera extensión de las desigualdades mundiales.

origen a la Tabla 11.2[5].

TABLA 11.2 *Rentas acumuladas*

Grupo de renta	*EE.UU 1980*	*EE.UU 1990*	*Holanda 1959*	*Holanda 1985*	*Mundial 1989*
20% inferior	*5,2*	*4,6*	*5,0*	*7,8*	*1,4*
40% inferior	*16,7*	*15,4*	*16,9*	*21,7*	*3,3*
60% inferior	*34,2*	*32,0*	*34,3*	*39,8*	*5,6*
80% inferior	*58,5*	*55,8*	*57,0*	*63,2*	*17,3*
100% inferior	*100,0*	*100,0*	*100,0*	*100,0*	*100,0*

La Figura 11.6 muestra dos de las curvas de Lorenz resultantes, halladas ajustando curvas lisas a los datos de la segunda y quinta columnas de la Tabla 11.2.

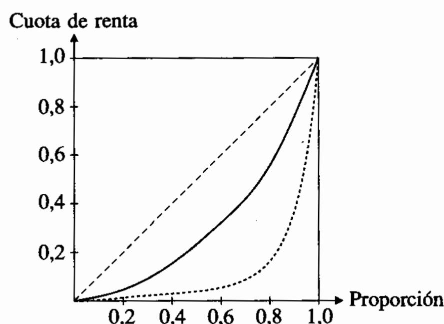


FIGURA 11.6 Curvas de Lorenz aproximadas para los EE.UU. en 1990 (curva sólida) y mundial en 1989 (curva punteada). La diagonal representa la igualdad perfecta.

La cuestión es ahora la siguiente: si la distribución de la renta viene descrita por la función continua de densidad $f(r)$, como en la Sección 10.4, ¿cómo se halla la curva de Lorenz? Para responder a esto tenemos primero que considerar la *función de distribución* (acumulada) $F(r)$ de la Sección 10.4, cuyo valor para cada nivel r de renta representa la proporción de la población que tiene rentas $\leq r$. Así el valor de esta función viene dado por la integral

$$F(r) = \int_0^r f(x)\,dx$$

que, evidentemente, satisface $F'(r) = f(r)$ para todo nivel r de renta. Supondremos que $f(r) > 0$ para todos los niveles de renta $r \geq 0$, lo que implica que $F(r)$ es estrictamente creciente. Además, suponiendo que todo el mundo tiene una renta, aun cuando sea muy pequeña, se debe verificar que $F(0) = 0$. También $F(\infty) = 1$, porque todo el mundo tiene renta finita, aun cuando algunos individuos puedan tener una renta extraordinariamente elevada. Aquí $F(\infty) = 1$ es una abreviatura de $F(r) \to 1$ cuando $r \to \infty$.

La variable sobre el eje horizontal de la gráfica de la curva de Lorenz es la proporción $p = F(r)$ de la población con rentas $\leq r$. La construcción de la curva de Lorenz requiere considerar la inversa de esta función, $r = R(p)$, que también es estrictamente creciente.

Se puede dar una interpretación importante de la función $R(p)$. Para cada $p \in [0, 1]$, el valor $R(p)$ es aquel nivel de renta para el que la proporción p de la población tiene una renta $r \leq R(p)$;

[5] Como ocurre a menudo con datos de este tipo, los errores de redondeo implican que las cifras de la Tabla 11.1 no suman el 100% en todos los casos.

por la definición de función inversa, $F(R(p)) = p$. Por ejemplo, cuando $p = 1/2$, el nivel de renta $R(1/2)$ tiene la propiedad de que la mitad de la población percibe una renta $r \leq R(1/2)$, mientras que la otra mitad percibe $r > R(1/2)$. Este nivel "medio" de renta se llama generalmente la *mediana* de la distribución. Ciertos intervalos de renta entre valores distintos de $R(p)$ reciben también nombres apropiados —por ejemplo el intervalo $[R(0{,}2), R(0{,}4)]$ se llama el *segundo quintil*, $[R(0{,}6), R(0{,}7)]$ el *séptimo decil*, y así sucesivamente. Los diferentes valores de $R(p)$ se llaman generalmente *percentiles* o *estadísticos de orden*.

Por la regla de derivación de la inversa de una función (ver (7.24), Sección 7.6), tenemos que

$$R'(p) = \frac{1}{F'(r)} = \frac{1}{f(r)} = \frac{1}{f(R(p))} \tag{11.14}$$

Esto vale para todo $p \in (0,1)$ porque hemos supuesto que $f(r) > 0$ para todos los niveles r renta.

La curva de Lorenz es la gráfica de la función $L(p)$ cuyo valor para cada p es la cuota de participación en la renta total que corresponde a la fracción p más pobre de la población. La renta total viene dada por $n\int_0^\infty r\, f(r)\, dr$, donde n es el número total de individuos de la población. Como $R(p)$ es el nivel de renta del más rico entre la fracción p más pobre de la población, la renta total de este grupo es $n\int_0^{R(p)} r\, f(r)\, dr$. Así tenemos

$$L(p) = \frac{n\int_0^{R(p)} r\, f(r)\, dr}{n\int_0^\infty r\, f(r)\, dr} = \frac{1}{m}\int_0^{R(p)} r\, f(r)\, dr \tag{11.15}$$

donde m es la renta media $\int_0^\infty r\, f(r)\, dr$. Como $0 \leq \int_0^{R(p)} rf(r)\, dr \leq \int_0^\infty rf(r)\, dr$, la ecuación (11.15) implica que $0 \leq L(p) \leq 1$ para todo $p \in [0,1]$. Se puede hallar la pendiente de la curva de Lorenz usando la regla de derivación (10.22), Sección 10.3. En efecto,

$$L'(p) = \frac{1}{m}R(p)f(R(p))R'(p) = \frac{R(p)}{m}$$

donde la segunda igualdad se deduce de (11.14). Así, la pendiente de la curva de Lorenz es igual a la razón entre el nivel $R(p)$ de renta y la renta media m. Esta pendiente crece paulatinamente desde $0 = R(0)$ cuando $p = 0$ hasta "$\infty = R(1)$" cuando $p = 1$. En particular, derivando una segunda vez se obtiene

$$L''(p) = \frac{R'(p)}{m} = \frac{1}{mf(R(p))} > 0$$

para todo $p \in (0,1)$, lo que implica que una curva de Lorenz es estrictamente convexa. Como muestra la Figura 11.6, cada curva de Lorenz tiene tangente horizontal en $p = 0$ y tangente vertical en $p = 1$. Finalmente, $L'(p) = 1$ en el único punto donde $R(p) = m$ luego para $p = F(m)$. Para $0 < p < F(m)$ se tiene que $L'(p) < 1$, luego la curva de Lorenz crece inicialmente más lentamente que la recta de 45° de inclinación. En $p = F(m)$ la distancia horizontal entre la curva de Lorenz y la recta de 45° de inclinación alcanza un máximo. Para $F(m) < p < 1$ se tiene $L'(p) > 1$, luego la curva de Lorenz termina creciendo más rápidamente que la recta hasta que las dos se vuelven a cortar en $p = 1$. Esto prueba, en particular, que $L(p) < p$ en el intervalo abierto $p \in (0,1)$.

Se puede usar también la curva de Lorenz para definir una medida común G de la desigualdad de renta, generalmente conocida como el **coeficiente de Gini**.[6] Geométricamente, G es el doble del área del conjunto que está limitado por la recta de 45° (la diagonal) y la curva de Lorenz. Este área

[6] El nombre viene del italiano Corrado Gini, quien la introdujo en 1912 y, aparentemente, descubrió la curva de Lorenz independientemente de Lorenz. Su definición era la integral doble $G = (1/2m)\int_0^\infty\int_0^\infty |r - r'|f(r)f(r')\, dr\, dr'$, pero es equivalente a la definición que damos aquí. No demostramos esta equivalencia porque no estudiamos integrales dobles en este libro.

se puede representar como la diferencia entre las integrales $\int_0^1 p\,dp = 1/2$ y $\int_0^1 L(p)\,dp$. Así

$$G = 2\left[\frac{1}{2} - \int_0^1 L(p)dp\right] = 1 - 2\int_0^1 L(p)\,dp \qquad \textbf{coeficiente de Gini} \qquad (11.16)$$

De aquí se deduce que $0 < G < 1$. Este coeficiente se aproxima a su extremo inferior $G = 0$ cuando la curva de Lorenz se acerca a la diagonal. Esto ocurre cuando la renta tiende a ser distribuida más equitativamente, con cada fracción más pobre p de la población aproximándose a percibir la cuota completa p de participación de la renta total disponible. El coeficiente se aproxima al otro extremo $G = 1$ cuando la curva de Lorenz se aleja de la diagonal. Esto ocurre cuando la renta se distribuye más desigualmente, con cada fracción p más pobre de la población aproximándose a la cuota cero de participación de la renta total disponible, y una pequeña fracción decreciente de personas muy ricas aproximándose a la percepción del total de renta disponible. Generalmente, conforme la curva de Lorenz baja, la distribución de la renta se hace más desigual y el coeficiente de Gini crece.

Problemas

1 Dibujar la curva de Lorenz para las columnas primera, tercera y cuarta de la Tabla 11.2.

2 Hallar los valores del coeficiente de Gini para las cinco distribuciones de la Tabla 11.2.